



MAGIC Meeting Minutes

December 2, 2015

Attendees

Jack Brassil	NSF
Emre Brooks	UT, San Antonio
Rich Carlson	DOE/SC
Jill Gemmill	Clemson
Dan Katz	NSF
Lavanya Ramakrishnan	LBL
Peter Lyster	NCO
Miron Livny	U. Wisconsin
Dave Martin	ANL
Grant Miller	NCO
JP Navarro	U Chicago, ANL
Greg Peterson	
Paul Redfern	Cornell
Alan Sill	TTU

Proceedings

The meeting was chaired by Rich Carlson, DOE and Dan Katz of the NSF. Jack Brassil of the NSF gave a presentation on the NSF FutureCloud program.

NSFFutureCloud: Jack Brassil

The NSF FutureCloud Program gave two \$10M awards in mid-2014 for CloudLab (PI-Ricci) with sites in Utah, Wisconsin, and Clemson and Chameleon (PI-Kate Keahy) with sites in Chicago and UT/Austin/TACC. The awards were to:

- Support multiple, large-scale concurrent short and long-term experiments.
- Involve novel computing architectures and applications
- Engage the research community as both users and developers

CloudLab provides a flexible, scalable infrastructure to enable exploration of fundamental science in the cloud built by and for the research community. CloudLab is a meta-cloud for building clouds enabling users to build their own clouds on CloudLab resources. It is agnostic to specific cloud software. It provides control and visibility all the way to bare metal. It is sliceable so it can support multiple experiments at once. CloudLab provides one facility, one account over three locations and includes 5,000 cores each (15,000 total), 8-16 cores per node, baseline 8GB RAM/core using the latest virtualization hardware. It provides 10Gb to nodes using SDN and 100 Gb to Internet2 and AL2S and enables partnerships with multiple vendors.

As of July 2015 there were 125 CloudLab projects with 440 users and 5600 experiments. They are exploring extreme and emerging cloud architectures, evaluating design choices for both hardware and software, studying geo-distributed data centers for low-latency, enabling real-time and near-realtime compute services and enabling data-intensive computing. Applications are exploring resource allocation, big data frameworks, traffic engineering, abstractions, next-generation stacks in clouds, anomaly detection, security, application-driven cloud architectures, and composing services.

FOR OFFICIAL GOVERNMENT USE ONLY

c/o National Coordination Office for Networking and Information Technology Research and Development

Suite II-405 · 4201 Wilson Boulevard · Arlington, Virginia 22230

Phone: (703) 292-4873 · Fax: (703) 292-9097 · Email: nco@nitrd.gov · Web site: www.nitrd.gov

CloudLab is federated with GENI so GENI accounts can be used and vice-versa. GENI provides 50 small clusters around the country and large clusters (100s of nodes) at several sites (ExoGENI). GENI provides wireless and mobile services and connects international partners.

CloudLab is open and in use. Hardware is being upgraded and refreshed regularly.

Chameleon is building a reconfigurable experimental testbed for cloud research. It is large scale to support big data, big compute, big instrument research with 650 nodes (14,500 cores), 5 PB disk storage over 2 sites connected by 100 Gbps networking. It is reconfigurable for users to configure resources uniquely for their research in reproducible ways. It is allied with GENI, Grid5000 and other experimental testbeds. Hardware includes:

- Multiple switches for 42 compute servers each with 128 GB RAM and 4 storage nodes X 2
- Chameleon core network of 100 Gbps
- Heterogeneous cloud units : to be available in 2016
- Core services for 3.6 PB central file systems and front end and data movers

The experimental work flow is being implemented to support experiment design, resource discovery, provisioning resources, configuring and interacting, monitoring, and analyzing to iteratively design the experiment and repeat the cycle.

Today there are over 500 users implementing 140 projects.

MAGIC member discussion

Miron Livny indicated that the Open System Grid (OSG) currently supports 3 million compute hours per day. How can cloud architectures contribute to the OSG and what would constitute a cloud? It was suggested that Miron Livny talk to Kate Keahy to explore ways to cooperate and interoperate.

CloudLab and Chameleon are meta-clouds that host cloud facilities. They are not in and of themselves a cloud.

Dan Katz indicated that the SI-2 program at NSF will explore software of interest to users and turn that software into infrastructure. Proposals for small develop-led projects are due in March 2016. Proposals for larger numbers of users are due in July-August. Proposals for SI-2 Software Institutes are being considered now.

DOE/SC is in flux now to respond to the NSCI initiative. DOE/SC will consider production with Exascale computing, streaming data, workflows, mining, simulation of workflows, and dynamic distributed research programs.

The NSF NeTS program is funding research on core networking research for primarily SDN and Wide-Area networks. NSF is also funding network infrastructures under GENI and US Ignite.

OSG is increasing its usage due to the increase in LHC data production. LIGO is an OSG user, using 200,000 hours per day. They are working on data provisioning in a complex resource environment. Management of OSG is focusing on traceability versus identity. The concept is to trust the VO rather than trusting the specific users.

Jill Gemmill: Clemson is a hosting site. Under FeduShare they are supporting federated logon with an SSH window from one campus cluster to another. They provide access to OSG and JetStream/XSEDE operating under user PKI certificates and campus logons.

Alan Sill: The Open Grid Forum (OGF) publishes documents for free as an informational community service for practices and standards.

Next MAGIC Meeting

January 6, 2016, 2:00-4:00 Eastern, NSF Room II-575